DOCUMENT RESUME

ED 272 291 PS 015 922

AUTHOR Davidson, Philip M.
TITLE Analyses of Interconcept Synchrony: Some Conceptual Issues and a Compromise Solution.
PUB DATE May 86
NOTE 25p.; Paper presented at the Biennial Meeting of the Midwestern Society for Research in Lifespan Development (Akron, OH, May, 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Development; *Cognitive Structures; *Criteria; Difficulty Level; Methods Research; *Research Methodology; Scoring; *Validity
IDENTIFIERS Connectivity; *Developmental Patterns; *Interconcept Consistency; Synchrony

ABSTRACT

Determining the extent of consistency between distinct concepts has become increasingly important in current theorizing about the development and organization of cognitive structure. This paper discusses several conceptual and measurement issues that must be addressed in analyses of interconcept consistency: the distinction between connectivity and synchrony, validity of measures, comparability of scoring criteria, and the distinction between intrinsic and extrinsic task difficulty. It is argued that the goal of equating task validity must often conflict with the goal of equating tasks for intrinsic difficulty. Therefore, an analytic method is called for which is sensitive to intertask asynchrony while controlling for differences in extrinsic difficulty. The use of a modified Wilcoxon procedure (Zwick et al., 1982) for this purpose is illustrated with data from a sample of 5- to 7-year-old subjects. An expected pattern of asynchrony is found, supporting the usefulness of the method. However, the main finding is a lack of asynchrony among a variety of quantitative and nonquantitative concepts during this age period. (Author/RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED272291

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

Analyses of Interconcept Synchrony:

Some Conceptual Issues and a Compromise Solution

Philip M. Davidson

Michigan State University

444 Erickson Hall

East Lansing, MI 48824

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Philip M. Davidson
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Biennial Meeting of the Midwestern Society for Research in Lifespan Development, Akron OH, May, 1986.

PS 015922

Abstract

Determining the extent of consistency between distinct concepts has become increasingly important in current theorizing about the development and organization of cognitive structure. This paper discusses several conceptual and measurement issues that must be addressed in analyses of interconcept consistency: the distinction between connectivity and synchrony, validity of measures, comparability of scoring criteria, and the distinction between intrinsic and extrinsic task difficulty. It is argued that the goal of equating task validity must often conflict with the goal of equating tasks for intrinsic difficulty. Therefore an analytic method is called for which is sensitive to intertask asynchrony while controlling for differences in extrinsic difficulty. The use of a modified Wilcoxon procedure (Zwick et al., 1982) for this purpose is illustrated with data from a sample of 5- to 7-year old subjects. An expected pattern of asynchrony is found, supporting the usefulness of the method. However, the main finding is a lack of asynchrony among a variety of quantitative and nonquantitative concepts during this age period.

Analyses of Interconcept Synchrony: Some Conceptual Issues and a Compromise Solution

One of the major theoretical debates about cognitive development has to do with the consistency among different conceptual achievements. For example, during the 1970's Flavell (e.g., 1982) and others argued against expecting the child's mind to be "homogenous," or typically at the same level of development across distinct conceptual domains. One source of evidence usually cited is correlational: some reports indicate that when age is controlled, correlations between distinct conceptual abilities are low (e.g., discussed by Pinard & Laurendeau, 1969). It is fair to say that recent years have seen an increasingly common assumption that development occurs within discrete, delimited domains, each of which progresses at an independent rate (e.g., Fischer, 1980).

In this paper I first review the measurement and definitional issues involved in determining relations between separate concepts, in cross-sectional data. I argue that our strategies for dealing with these issues will often be in conflict. This means there can be no simple statistical formula for definitively establishing or disconfirming interconcept consistency. I then present one strategy for analyzing cross-sectional data that reaches a compromise between conflicting analytical objectives. The analyses are illustrated with data from a sample of 72 children who were tested on eight

different types of conceptual abilities (listed in Table 1).

Insert Table 1 about here

First of all, in considering consistency in development, we need to distinguish between interconcept connectivity and interconcept synchrony. By connectivity I mean the claim that one conceptual domain may have an effect on the development of another conceptual domain. By synchrony I mean the claim that certain distinct concepts are typically at the same level of development at a given time.

These two aspects of consistency are in principle independent. For example, two concepts could be developmentally connected--but not at the same level--if one is a prerequisite for the other. On the other hand, several concepts could typically be at the same level--but not necessarily connected-- if there are constraints built into the nature of conceptual development itself, so that development occurs at the same pace, regardless of domain. This is a basic assumption in Case's (1985) model, for instance.

Measurement Issues

There are three basic measurement issues we need to keep in mind. First, are the separate tasks equally valid measures of the concepts they are intended to measure? This is essentially a conceptual issue having to do with designing tasks that can be

solved using the specific (hypothesized) cognitive operations we are interested in studying. There are ways to test task validity empirically, but the method used would depend on a conceptual analysis of a particular task and its relation to a particular theory. For example, Baroody (1982) demonstrates that a particular task is not solved easier by adults than by children. This is persuasive evidence against this task's validity because the conceptual framework on which the task is based makes strong developmental assumptions. One can easily imagine other experimental situations (for instance, studies of discrimination between content domains; cf. Turiel, 1983) in which lack of development from childhood to adulthood is not an appropriate test of task validity, given the theoretical framework within which the task is proposed.

Second, there is the issue of equating the appropriateness of the scoring criteria. The problem here is deciding what counts as sufficient evidence of the concepts in question. For example, children may solve an addition problem by using manipulable objects, or by mental arithmetic. As long as we are interested only in arithmetic concepts, we might choose either of these responses as adequate, depending on the particular ability (or level of ability) we are studying. However, if we are interested in <u>comparing</u> arithmetic concepts with other concepts, we need a scoring metric that is equally applicable to various types of concept acquisition,

rather than one that is derived from properties of specific tasks.

In the present research I adapted a scoring metric that has been employed by Piaget (e.g., Inhelder & Piaget, 1964) in a variety of conceptual domains. The system, which is summarized in Table 2, has six levels: First, there are two levels reflecting failure on a task, followed by two levels of success through trial-and-error, and two levels of success through deductive strategies. In other words, the scoring procedure starts with deductive strategies as a principled definition of mastery that cuts across conceptual domains. Tasks in separate domains are then analyzed to determine particular response behaviors that constitute deductive solutions, and particular behaviors that constitute trial-and-error. The interrater reliability coefficients for this coding system are listed in Table 1. The reliabilities are uniformly high, indicating that this method can be applied to a variety of tasks, some of which are classical and some of which are original to this project.

Insert Table 2 about here

The third issue to keep in mind is equating the difficulty of the different tasks. The problem here is that virtually all measures of concept formation entail psychological processes other than the conceptual abilities we are primarily interested in. To give some simple examples: drawing tasks involve a strong element

of motor skill, even though what we are interested in may be perspectivism. Some tasks can be administered with little reliance on language processes, but others, such as class inclusion, are essentially verbal. Some tasks, such as conservation problems, depend on placing perceptual processes in conflict with conceptual processes. In others, such as seriation problems, perceptual processes and conceptual processes are mutually supportive.

My point is that, at least in many cases, the goal of measuring distinct concepts with equal validity is in conflict with the goal of providing tasks that are of equal difficulty. In order to be a valid measure of concept A, a task may require extensive verbal instructions, whereas a valid measure of concept B may not. In other words, the normal state of affairs is that measures of different abilities may vary in difficulty for legitimate reasons that nevertheless have little to do with the intrinsic difficulty, or developmental status, of the cognitive competency we are interested in. Variations between measures of concept formation which have to do with memory, perception, motor skill, verbal comprehension, and so forth can be described as extrinsic sources of task difficulty (in so far as such variables are not called on as explanatory constructs in the relevant theory, here taken for expository purposes to be a theory of concept formation).

The present distinction between intrinsic and extrinsic factors should not be confused with the distinction between competence and

performance in cognitive development (*e.g.*, Flavell & Wohlwill, 1969). A theory which takes the competence-performance distinction seriously would surely aim to specify which performance variables are intrinsic to the cognitive ability in question, and which are incidental to a particular measure of that ability. For example, ability to draw a straight line might be considered as intrinsically related to artistic competence, whereas it might be considered as incidental to a particular measure of geometrical problem solving. Thus, extrinsic factors are those due to the nature of specific tasks, and which vary when different tasks are used to assess the same competence (*e.g.*, use of verbal reasoning rather than pencil and paper construction, in the assessment of geometrical problem solving).

Analytic Strategies

Let us consider the two basic analytic strategies for investigating connectivity and synchrony between two variables. First, there are the norm-referenced correlational approaches, which compute the association between standardized or rank-order scores. The correlations between variables in the present study are found in Table 3; zero-order correlations are shown above the diagonal, and age-partialed correlations below the diagonal. In this example, all the age-partialed correlations are significant at the .01 level. Assuming valid measures and valid scoring criteria, correlations can

be reasonable evidence for connectivity. However, relations found between variables are essentially meaningless without a theoretical specification of what ought to be connected, and why. Within the correlational approach, such specifications can be evaluated in increasing detail through factor analysis, multiple regression techniques, and causal modeling.

Insert Table 3 about here

One merit of correlations is that they control for the unavoidable extrinsic differences in task difficulty just mentioned, since they are based on (norm-referenced) standarized or ordinal scores. However, this makes them useless for investigating interconcept synchrony. To assess synchrony, discrete criterion-referenced approaches are typically used to compute the probability that subjects who attain a specific response criterion on Concept A also attain the same (or higher or lower) criterion on Concept B.

This kind of analysis is most appropriate when we are reasonably assured that extrinsic demands (having to do with memory, perception, motor skill, verbal comprehension, etc.) are equated between the two measures. As an example, elsewhere (Davidson, in press) I report on two types of problems having to do with class logic. The stimulus materials, the form of verbal questioning, and other task parameters are identical for both problem types -- the

only difference being whether the questions refer to overlapping or non-overlapping classes. A significant asynchrony is found, with reasoning about overlapping classes lagging behind reasoning about non-overlapping classes. In this case, I would argue, it is appropriate to make direct comparisons of raw scores to determine which of these two types of reasoning is intrinsically more difficult.

However, in the general case it is not possible to experimentally control all extrinsic sources of task difficulty without affecting task validity. Consequently, it is quite likely that many findings of developmental sequences (asynchronies) among abilities are due to analyses that unintentionally capitalize on systematic intertask differences in level of extrinsic difficulty. Therefore, it would be well to have some statistical means of equating these effects, while also taking into account information about the synchrony/asynchrony between pairs of abilities. The approach I will describe is a hybrid of the two mentioned above: In common with correlations, the scores used are norm-referenced; this controls extrinsic sources of difficulty. Unlike correlations, the analysis computes the degree and direction of any asynchrony between pairs of variables.

A Test for Interconcept Asynchrony

The first step of the procedure in the present example is to divide the sample into thirds with respect to each variable,

representing low, average, and high performance. That is, each subject is assigned a score of low, average, or high, relative to the sample, on each variable. (Note that any number of subdivisions could be used; grouping by thirds is convenient for the present N.) Pairs of variables are then crosstabulated, to determine the frequency of asynchronous scores. Table 4(a) shows the example of number conservation scores crosstabulated with seriation scores. The main diagonal represents the number of subjects with equal scores on both variables ($n = 35$). The upper left part of the table indicates that 4 subjects have conservation scores 2 levels above seriation scores, and 19 have conservation scores 1 level above seriation. The lower right part of the table shows 3 subjects with conservation scores 2 levels below seriation, and 11 subjects with conservation 1 level below seriation.

Insert Table 4 about here

Notice that there is a lot of asynchrony in the table; fewer than half of the subjects have conservation scores exactly corresponding to seriation. This illustrates that the assignment of norm-referenced scores does not reduce information about asynchrony contained in the data. The null hypothesis in the present analysis is that the extent of asynchrony with conservation in advance is equal to the extent of asynchrony with seriation in advance. This

hypothesis can be tested by a modified Wilcoxon procedure described by Zwick, Neuhoff, Marascuilo, and Levin (1982). The computation results in a Z score indicating the extent to which the given distribution (e.g., Table 4(b)) departs from the normal distribution expected when the amount of asynchrony in one direction equals that in the other direction. In this example, the procedure yields a Z-approximation of 1.36, and the null hypothesis is supported.

If we repeat this analysis with all pairs of variables in the data set, we obtain the values shown in Table 5. Negative Z scores indicate that the row variable lags behind the column variable; for instance, in the upper left corner of the table, arithmetic ability lags behind number conservation (Z = -1.07, ns.). Positive values indicate the row variable is in advance of the column variable. None of the pairwise asynchronies is significant. It is important to realize that this does not mean that children are usually at the same level on each pair of variables; it means that the number who are advanced on one variable is not significantly different from the number who are advanced on the other.

Insert Table 5 about here

It is also of interest to ask whether these results would be different if age were controlled. Using analyses based on frequency tabulations, it is possible to (approximately) control age by

considering age groups separately. Table 6 shows the results when the younger half of the sample (n = 35, mean age = 5 years 8 months) is analyzed separately from the older half (n = 37, mean age = 6 years 10 months). Each cell above the diagonal contains two Z scores: the upper value is for the younger group and the lower value is for the older group.

Insert Table 6 about here

Again, the great majority of comparisons show little asynchrony for either the younger or the older group. However, there is a pattern having to do with arithmetic performance: Younger children show arithmetic lagging behind all other variables, whereas older children show the opposite pattern. It is possible to test this shift in the direction of asynchrony, using a two-sample modified Wilcoxon test (Zwick et al., 1982). The null hypothesis in this case is that the direction of asynchrony is the same for both groups. The values below the diagonal of Table 6 show the results of these 2-sample contrasts, and confirm that the shift in direction of asynchrony for arithmetic is significant with respect to most other variables.

This finding is not too difficult to interpret. The older group has received more formal schooling than the younger group, and of all the variables in this study, arithmetic is the one that is most directly influenced by formal instruction. Therefore, these

analyses have detected exactly the pattern of asynchrony we should expect: younger children have a deficit in arithmetic relative to other abilities, and older children have an overall advantage in arithmetic relative to other abilities.

Conclusion

In summary, I have argued that in analyzing interconcept consistency in cross-sectional data, first of all it is necessary to have a principled scoring metric that is applicable across the range of concepts to be studied. Second, in analyzing interconcept asynchrony, I have proposed that it is prudent to avoid capitalizing on unavoidable inter-task differences in extrinsic sources of difficulty. This calls for an analytic strategy that is in some ways a compromise between strictly norm-referenced analyses and strictly criterion-referenced analyses. I have suggested one such compromise approach. The analyses shown in Table 6 indicate that this approach is in fact sensitive to the kinds of asynchrony that we should expect it to be sensitive to.

Nevertheless, the main finding from these analyses is that synchrony among a variety of quantitative and nonquantitative concepts is the rule rather than the exception. Together with the correlational evidence, this finding is consistent with the construct of general developmental stages in cognitive development. However, as noted previously, it is well to bear in mind that

concepts do not necessarily develop in synchrony because they are extensively interconnected. My working hypothesis is that concepts develop within distinct conceptual domains (cf., Turiel & Davidson, 1986), but they do develop in synchrony because the intrinsic mechanisms of concept formation within separate domains are similar in fundamental ways.

References

Baroody, A. (1982). Evaluating the validity of Brainerd's cardinality task. Child Study Journal, 12, 79-87.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

Davidson, P. (in press). Developing coordinations of class unions, intersections, and complements. Cognitive Development.

Fischer, K. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. Psychological Review, 87, 477-531.

Flavell, J. (1982). On cognitive development. Child Development, 53, 1-10.

Flavell, J., & Wohlwill, J. (1969). Formal and functional aspects of cognitive development. In D. Elkind & J. Flavell (Eds.), Studies in cognitive development (pp. 67-120). New York: Oxford University Press.

Inhelder, B., & Piaget, P. (1964). The early growth of logic in the child. London: Routledge & Kegan Paul.

Pinard, A., & Laurendeau, M. (1969). "Stage" in Piaget's cognitive-developmental theory: Exegesis of a concept. In D. Elkind, & J. Flavell (Eds.). Studies in cognitive development (pp. 121-170). New York: Oxford University Press.

Turiel, E. (1983). The development of social knowledge. New York: Cambridge University Press.

Turiel, E. & Davidson, P. (1986) Heterogeneity, inconsistency, and asynchrony in the development of cognitive structures. In I. Levin (Ed.), Stage and structure: Reopening the debate (pp. 106-143). Norwood, NJ: Ablex.

Zwick, R., Newhoff, V., Marascuilo, L., & Levin, J. (1982). Statistical tests for correlated proportions: Some extensions. Psychological Bulletin, 92, 258-271.

Table 1

Summary of Composite Performance Variables

Conceptual Domain	Performance variable	Coding reliability $r(22)$
Number	1. Arithmetic operations (+,-,x,÷)	.98
	2. Number conservation	.96
Logic	3. Class inclusion	.99
	4. Class vicariance	.98
	5. Seriation	.97
Functions	6. Displacement (functions as movement from one location to another)	.98
	7. Exchange (functions as exchange of one object for another)	.96
	8. Morphisms (functions as correspondences that preserve an operation)	.99

Note. Subjects (N = 72) ranged in age from 5 years 0 months to 7 years 10 months. They were tested on several problems of each conceptual ability. Scores on each set of problems were then composited to yield the eight performance variables listed.

Coding reliabilities were obtained by recoding one-third (n = 24) of the sample by a second rater.

Table 2.

Summary of Scoring Criteria

Score	Brief Description
Level 0.	No success; no evidence of understanding how to approach the problem; gives no answer, or answers appear to be wild guesses.
Level 1.	No success; some understanding of the problem; makes reasonable efforts; answers appear to be the product of reasoning rather than guessing.
Level 2.	Partial success by groping. Examples: Exchange and displacement functions: Majority of incorrect responses, all by trial-and-error. Number conservation: Inconsistent response: correct on one question, incorrect on the other; no adequate explanation. Arithmetic: Incorrect answer; finds correct answer after trial-and-error using concrete objects. Class inclusion: Compromise solution; e.g., avoids asserting that the part is greater than the whole by stating they are the same. Vicariance: Incorrect answer after vacillating and considering the correct answer. Seriation: Compromise solution; e.g., constructs a V-shaped series by trial-and-error.
Level 3.	Substantial or complete success by groping. Examples: Exchange and displacement functions: Correct responses, all by trial-and-error; some preceded by vacillating, incorrect responses. Number conservation: Both answers correct, both supported by irrelevant justifications (e.g., counting).

Table 2 (Cont'd)

Score	Brief description
	Arithmetic: Finds correct answer directly, but relying on concrete objects.
	Class inclusion and vicariance: Correct answer after initial incorrect answer, vacillation, repetitions of question.
	Seriation: Correct series constructed after extensive trial-and-error.
Level 4.	Complete success but lacking evidence of thorough mastery of the task. For instance, succeeds through a mixture of anticipatory and trial-and-error strategies; or, can explain reasoning on one problem, but not on another.
Level 5.	Complete success, and shows a thorough grasp of the task. Examples:
	Exchange and displacement functions: Correct answers to all problems without trial-and-error.
	Number conservation: Correct answers, resists countersuggestions, gives appropriate explanations (e.g., compensation).
	Arithmetic: Correct answer given immediately, without any need for concrete materials.
	Class inclusion and vicariance: Correct answer without initial vacillation, resists countersuggestion, explains reasoning.
	Seriation: Correct construction of series, including intercalating additional pieces, by systematic anticipation.

Table 3.

Product-moment correlations among performance variables (Partial correlations controling age are below the diagonal)

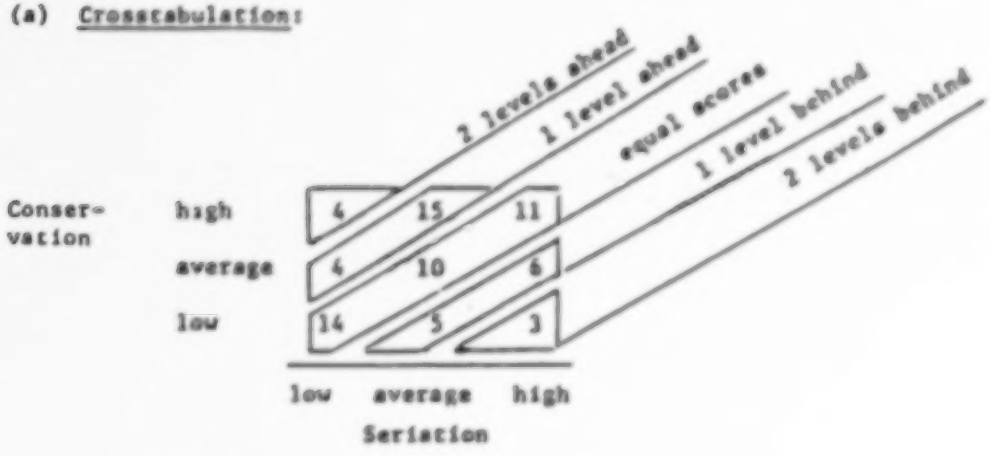
Performance variable	1.	2.	3.	4.	5.	6.	7.	8.
1. Arithmetic	--	.46	.55	.61	.60	.74	.67	.43
2. Number conservation	.41	--	.37	.45	.36	.61	.40	.37
3. Class inclusion	.49	.34	--	.48	.31	.57	.37	.45
4. Vicariance	.53	.41	.42	--	.48	.53	.62	.56
5. Seriation	.48	.30	.44	.39	--	.49	.61	.41
6. Displacement funct's.	.64	.39	.31	.44	.37	--	.68	.44
7. Exchange functions	.56	.34	.31	.33	.32	.38	--	.53
8. Morphisms	.35	.42	.43	.54	.38	.41	.32	--

Note. All correlations are significant at $p < .01$; correlations of .36 or greater are significant at $p < .001$.

Table 4.

Illustration of asynchrony analysis: Number conservation vs. seriation.

(a) Crosstabulation:



(b) Difference scores:

	Conservation (with respect to seriation)				
	2 levels ahead	1 level ahead	Equal scores	1 level behind	2 levels behind
n =	4	19	35	11	3

(c) Modified Wilcoxon analysis:[a]

$T(+) = 1,550.5$, $Z = 1.36$, $p = .17$ (2-tailed).

[a]Described by Zwick, Neuhoff, Marascuilo, & Levin, Psychological Bulletin, 1982, 92, 258-271.

Table 5

Z-Approximations for Pairwise Asynchrony Analyses (Whole Sample)

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. Arithmetic	--	-1.07	-.23	-.16	.56	-.48	-.36	.01
2. Conservation		--	.98	1.13	1.36	.93	1.08	.69
3. Class inclusion			--	.34	.46	-.20	.37	.01
4. Class vicariance				--	.20	-.58	.01	-.31
5. Seriation					--	-.85	.00	-.29
6. Displacement functions						--	.93	.01
7. Exchange functions							--	-.30
8. Morphisms								--

Note. Negative Z scores indicate that the row variable (i.e., the label at the left) lags behind the row variable (label at top of table). Positive values indicate that the column variable lags behind the row variable.

Table 6.

Normal approximations for Wilcoxon Tests Contrasting Pairs of Variables: One-Sample Tests for Younger and Older Groups above the Diagonal, Two-Sample Tests for Younger versus Older Groups below the Diagonal

Variable	Arithmetic	Conservation	Inclusion	Vicariance	Seriation	Disp. functions	Exch. functions	Morphisms
Arithmetic	--	-2.30**	-1.77*	-1.32	-.91	-1.19	-1.14	-2.03**
		.84	1.72*	1.67*	1.77*	.33	1.75*	1.94*
Conservation	2.51**	--	.81	.86	1.57	1.63	.91	.07
			.34	.73	.39	.34	.60	.89
Inclusion	2.79***	-.13	--	.21	.52	.63	.48	-.36
				.29	.12	-1.02	.00	.43
Vicariance	2.27**	-.10	.02	--	.43	.29	.23	-1.10
					-.16	-1.09	-.13	.53
Seriation	1.95*	-.82	-.27	-.46	--	-.14	-.22	-1.18
						-.95	.23	.64
Displacement functions	1.43	-1.64	-1.26	-1.04	-.66	--	.00	-1.29
							1.38	1.44
Exchange functions	2.19**	-.23	-.30	-.25	.33	1.05	--	-.98
								.48
Morphisms	-2.88***	-.66	-.58	-1.15	-1.32	-2.07**	-1.06	--

Note. For one-sample tests, upper Z values are for the younger group (mean age = 5;8, n = 35) and lower Z values are for the older group (mean age = 6;10, n = 37). Negative values indicate row variable lags behind column variable; positive values indicate column variable lags behind row variable. Two-sample tests may be interpreted by comparing Z values for younger and older groups in corresponding one-sample tests.

*$p < .10$; **$p < .05$; ***$p < .01$ (two-tailed).